Mirror-Based Extrinsic Camera Calibration

Joel A. Hesch, Anastasios I. Mourikis, and Stergios I. Roumeliotis

Abstract This paper presents a method for determining the six degree-of-freedom transformation between a camera and a base frame of interest. A planar mirror is maneuvered so as to allow the camera to observe the environment from several viewing angles. Points, whose coordinates in the base frame are known, are observed by the camera via their reflections in the mirror. Exploiting these measurements, we determine the camera-to-base transformation analytically, without assuming prior knowledge of the mirror motion or placement with respect to the camera. The computed solution is refined using a maximum-likelihood estimator, to obtain high-accuracy estimates of the camera-to-base transformation and the mirror configuration for each image. We validate the accuracy and correctness of our method with simulations and real-world experiments.

1 Introduction

Cameras are utilized in a wide variety of applications ranging from surveillance and crowd monitoring, to vision-based robot localization. In order to obtain meaningful geometric information from a camera, two calibration procedures must be completed. The first is intrinsic calibration, that is, determining the internal camera parameters (e.g., focal length, principal point, and skew coefficients), which affect the image measurements. The second is extrinsic calibration, which is the process of computing the transformation between the camera and a base frame of reference. In a surveillance application, the base frame may be the room or building coordinate system, whereas on a mobile robot, the base frame could be the robot-body frame. Several authors have addressed extrinsic calibration of a cam-

Joel A. Hesch and Stergios I. Roumeliotis
University of Minnesota, Minneapolis, MN 55455, USA, e-mail: {joel|stergios}@cs.umn.edu

Anastasios I. Mourikis
University of California, Riverside, CA 92521, USA, e-mail: mourikis@ee.ucr.edu

era to another sensor (e.g., for odometry-to-camera [13], inertial measurement unit (IMU)-to-camera [14], or laser scanner-to-camera [1, 21]). These exploit measurements from both sensors to determine their mutual transformation. However, very little attention has been devoted to determining the *camera-to-base* transformation, for a generic base frame.

In this paper, we deal exclusively with extrinsic camera calibration. Our objective is to determine the camera-to-base transformation from observations of points whose coordinates in the base frame are known. We consider the most limiting case, in which the known points do not lie within the camera's field of view but can only be observed using a planar mirror. We maneuver the mirror in front of the camera to provide multiple views of the points. In our formulation, no prior information about the mirror motion or placement with respect to the camera is assumed. The configuration of the mirror and the camera-to-base transformation are treated as unknowns to be computed from the observations. The main contribution of this paper is an algorithm for determining the camera-to-base transformation *analytically*, which requires a minimum of 3 non-collinear points tracked in 3 images.

A direct approach to extrinsic camera calibration is to utilize all of the measurements in a maximum-likelihood estimator (MLE) for computing the unknown transformation [7]. This takes the form of a nonlinear least-squares problem, which seeks to iteratively minimize a nonconvex function of the unknown variables. While appealing for its ease of implementation, this method has two drawbacks. First, without an accurate initial guess, the minimization process may take several iterations to converge, or even fail to find the correct solution. Second, the MLE provides no framework for studying the minimal measurement conditions required to compute a solution. To address the first issue, in the method presented in this paper we first determine the transformation analytically, and then employ an MLE to refine the computed solution (cf. Section 3). Moreover, we determine the minimal number of measurements required for a unique solution. Finally, in Appendix 2 we comment on the extension of this work to robot-body 3D reconstruction.

2 Related Work

Before presenting our method, we first review the related work, which falls into two categories: (i) hand-eye calibration, and (ii) catadioptric systems. Hand-eye calibration is the process of determining the six degree-of-freedom (6 d.o.f.) transformation between a camera and a tool, which are both mounted on a robot manipulator [20, 2]. The hand-eye problem is solved by correlating the measurements of the camera and the encoders, which measure displacements of the robot joints. This process determines the pose of the camera with respect to the robot base. Subsequently, the camera-to-tool transformation is calculated by combining the estimated camera-to-robot-base transformation, and the robot-base-to-tool transformation, which is assumed to be known. This necessitates the availability of precise technical drawings, and limits the applicability of these methods, since they cannot determine the camera-to-base transformation for a generic base frame.

Next, we turn our attention to catadioptric systems, which are employed to perform synthetic multiple-view vision. Methods have been presented utilizing a single camera and planar [3, 8], or conic mirrors [9]. Others accomplish stereo vision with reflections from free-form surfaces [22], and a trinocular mirror-based vision system also exists [17]. Additionally, stereo is achieved with a static camera and a moving mirror [11], or with a moving camera and two stationary spherical mirrors of known radii [15]. While the use of mirror reflections relates these approaches to our work, the key difference is that we do *not* perform synthetic stereo, i.e., only a single observation of each point is available in each image.

Jang *et al.* demonstrate a system for 3D scene reconstruction using a moving planar mirror [10]. Exploiting a combination of fiducial points on the mirror and vanishing points in the reflections, they solve for the position of the mirror with respect to the camera. The 3D scene is determined based on synthetic stereo from multiple reflections. In contrast to this approach, we do not assume that the dimensions of the mirror, or its position with respect to the camera, are available. Finally, Kumar *et al.* determine the transformations between multiple cameras with non-overlapping fields of view, using mirror reflections of a calibration grid [12]. They require 5 views (per camera) of the calibration pattern to form a set of linear constraints, which are solved for the unknown transformations. In contrast to [12], our method requires only 3 images, each containing observations of 3 known points, to determine the camera-to-base transformation analytically.

3 Computing the Transformation

In this section, we describe our approach for analytically determining the transformation between the camera frame, $\{C\}$, and a frame of interest, $\{B\}$, from observations of 3 points whose coordinates in $\{B\}$ are known. Frame $\{B\}$ is arbitrary and without loss of generality, we will refer to $\{B\}$ as the "base frame." Example base frames vary by application, and may include: (i) the robot-body frame, if the camera is mounted on a robot, (ii) the room or building frame, if the camera is utilized in a surveillance application, and (iii) the rig mount, if the camera is part of a stereo pair.

We address the most limiting scenario in which the points are only visible through reflections in a planar mirror that is moved in front of the camera to provide multiple views of the scene. We exploit these observations to compute the transformation between $\{B\}$ and $\{C\}$, without knowledge of the mirror's placement or motion with respect to the camera (cf. Algorithm 1). In what follows, we present the measurement model and discuss its relation with the three-point pose estimation problem (P3P). We comment cases where a unique solution does not exist, and present an analytical method to compute the unknown transformation from a minimum of 3 points observed in 3 images that differ by rotations about two axes.

Algorithm 1 Computing the Camera-to-Base Transformation

```
Input: Observations of 3 points tracked in $N_c$ images
Output: Camera-to-base transformation $\{{}^{C}_{B}\mathbf{R}, {}^{C}\mathbf{p}_{B}\}$
  for each image in $N_c$ do
    Convert to three-point pose estimation problem (P3P)
    Solve P3P to obtain combined homogeneous/reflection transformation: $\{\mathbf{A}, \mathbf{b}\}_k$
  end for
  for each triplet of solutions $\{\mathbf{A}, \mathbf{b}\}_k$, $\{\mathbf{A}, \mathbf{b}\}_{k'}$, $\{\mathbf{A}, \mathbf{b}\}_{k''}$ do
    Compute mirror configurations from (14)
    Compute camera-to-base rotation ${}^{C}_{B}\mathbf{R}$ from (23)
    Compute camera-to-base translation ${}^{C}\mathbf{p}_{B}$ from (16)
  end for
  Utilize clustering to select the correct solution $\{{}^{C}_{B}\mathbf{R}, {}^{C}\mathbf{p}_{B}\}$
  Refine the solution using a maximum-likelihood estimator
```

Lastly, we summarize a maximum-likelihood approach for refining the computed transformation, a detailed discussion of which is available in [7].

3.1 Measurement Model

First, we present the measurement model that describes each of the camera observations. To simplify the presentation, in this section we focus on the case of a single point, observed in a single image. Consider a point $\mathbf{p}$, whose position with respect to frame $\{B\}$, ${}^{B}\mathbf{p}$, is known[1]. We seek to express the point $\mathbf{p}$ in the camera reference frame $\{C\}$. From geometry (cf. Fig. 1) we have two constraint equations:

$$ {}^{C}\mathbf{p}' = {}^{C}\mathbf{p} + 2 d_p {}^{C}\mathbf{n} \tag{1} $$

$$ d_p = d - {}^{C}\mathbf{n}^{\mathrm{T}} {}^{C}\mathbf{p} \tag{2} $$

where ${}^{C}\mathbf{p}'$ is the reflection of ${}^{C}\mathbf{p}$, ${}^{C}\mathbf{n}$ is the mirror normal vector expressed in the camera frame, d is the distance between the mirror and the camera, and d_p is the distance between the mirror and the known point (both distances are defined along the mirror normal vector). Note also that

$$ {}^{C}\mathbf{p} = {}^{C}_{B}\mathbf{R} {}^{B}\mathbf{p} + {}^{C}\mathbf{p}_{B} \tag{3} $$

where ${}^{C}_{B}\mathbf{R}$ is the matrix which rotates vectors between frames $\{B\}$ and $\{C\}$, and ${}^{C}\mathbf{p}_{B}$ is the origin of $\{B\}$ with respect to $\{C\}$. We substitute (2) and (3) into (1), and rearrange the terms to obtain:

[1] Throughout this paper, ${}^{X}\mathbf{y}$ denotes a vector $\mathbf{y}$ expressed with respect to frame $\{X\}$, ${}^{X}_{W}\mathbf{R}$ is the rotation matrix rotating vectors from frame $\{W\}$ to $\{X\}$, and ${}^{X}\mathbf{p}_{W}$ is the origin of $\{W\}$, expressed with respect to $\{X\}$. $\mathbf{I}_n$ is the $n \times n$ identity matrix, and $\mathbf{0}_{m \times n}$ is the $m \times n$ matrix of zeros.

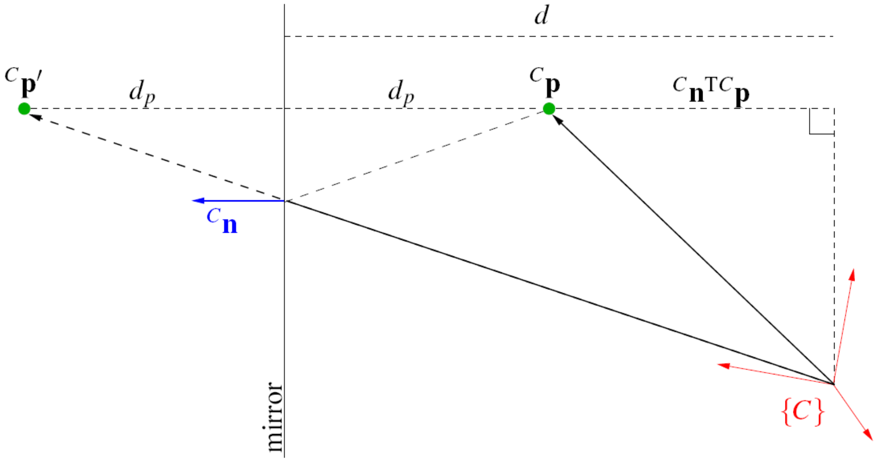


Fig. 1: Observation of the point ${}^{C}\mathbf{p}'$ which is the reflection of ${}^{C}\mathbf{p}$. In this figure, the mirror plane is perpendicular to the page. Only the reflected point is in the camera's field of view; the real point is not observed directly by the camera.

$$\begin{aligned} {}^{C}\mathbf{p}' &= \left(\mathbf{I}_3 - 2{}^{C}\mathbf{n}{}^{C}\mathbf{n}^{\mathrm{T}}\right){}^{C}\mathbf{p} + 2d\,{}^{C}\mathbf{n} \\ &= \left(\mathbf{I}_3 - 2{}^{C}\mathbf{n}{}^{C}\mathbf{n}^{\mathrm{T}}\right)\left({}^{C}_{B}\mathbf{R}{}^{B}\mathbf{p} + {}^{C}\mathbf{p}_B\right) + 2d\,{}^{C}\mathbf{n} \end{aligned} \tag{4}$$

which can be written in homogeneous coordinates as:

$$\begin{bmatrix} {}^{C}\mathbf{p}' \\ 1 \end{bmatrix} = \begin{bmatrix} \left(\mathbf{I}_3 - 2{}^{C}\mathbf{n}{}^{C}\mathbf{n}^{\mathrm{T}}\right) & 2d\,{}^{C}\mathbf{n} \\ \mathbf{0}_{1\times 3} & 1 \end{bmatrix} \begin{bmatrix} {}^{C}_{B}\mathbf{R} & {}^{C}\mathbf{p}_B \\ \mathbf{0}_{1\times 3} & 1 \end{bmatrix} \begin{bmatrix} {}^{B}\mathbf{p} \\ 1 \end{bmatrix}. \tag{5}$$

The reflection of $\mathbf{p}$ is observed by the camera, and this measurement is described by the perspective projection model:

$$\mathbf{z} = \frac{1}{p_3}\begin{bmatrix} p_1 \\ p_2 \end{bmatrix} + \eta = \mathbf{h}({}^{C}\mathbf{p}') + \eta, \quad {}^{C}\mathbf{p}' = \begin{bmatrix} p_1 & p_2 & p_3 \end{bmatrix}^{\mathrm{T}} \tag{6}$$

where η is the pixel noise, assumed to be a zero-mean, white Gaussian process with covariance matrix $\sigma_\eta^2\mathbf{I}_2$. Equations (4) and (6) define the measurement model, which expresses the observed image coordinates, $\mathbf{z}$, of the point as a function of the *known* position vector, ${}^{B}\mathbf{p}$, the *unknown* camera-to-base transformation, $\{{}^{C}_{B}\mathbf{R}, {}^{C}\mathbf{p}_B\}$, and the *unknown* configuration of the mirror with respect to the camera, $\{{}^{C}\mathbf{n},\ d\}$. Note that the transformation between the mirror and camera has 6 d.o.f., however, only 3 d.o.f. appear in the measurement equation. These are expressed by the vector $d\,{}^{C}\mathbf{n}$, which has 2 d.o.f. from the mirror normal, ${}^{C}\mathbf{n}$, and 1 d.o.f. from the camera-to-mirror distance, d. The remaining 3 d.o.f., which correspond to rotations about ${}^{C}\mathbf{n}$ and translations of the mirror-frame origin in the mirror plane, do not affect the measurements.

3.2 Three-Point Perspective Pose Estimation Problem

We now briefly review the three-point perspective pose estimation problem (P3P) and discuss how it relates to our problem. The goal of P3P is to determine the 6 d.o.f. transformation, $\{{}^C_B\mathbf{R}, {}^C\mathbf{p}_B\}$, between a camera frame, $\{C\}$, and a base frame, $\{B\}$, given the known coordinates of 3 non-collinear points, ${}^B\mathbf{p}_i$, $i = 1\ldots3$, in $\{B\}$, and their corresponding perspective projections, $\mathbf{z}_i$, in $\{C\}$, defined as:[2]

$$\mathbf{z}_i = \frac{1}{p_{3i}} \begin{bmatrix} p_{1i} \\ p_{2i} \end{bmatrix}, \quad {}^C\mathbf{p}'_i = \begin{bmatrix} p_{1i} & p_{2i} & p_{3i} \end{bmatrix}^{\mathrm{T}} \tag{7}$$

$$\begin{bmatrix} {}^C\mathbf{p}'_i \\ 1 \end{bmatrix} = \begin{bmatrix} {}^C_B\mathbf{R} & {}^C\mathbf{p}_B \\ \mathbf{0}_{1\times3} & 1 \end{bmatrix} \begin{bmatrix} {}^B\mathbf{p}_i \\ 1 \end{bmatrix}. \tag{8}$$

This problem has up to 4 pairs of solutions, where for each pair, there is one solution lying in front of the center of perspectivity and one behind it [4].

Equation (8) differs from (5) in that the former expresses a homogeneous transformation, while the latter describes a homogeneous transformation followed by a reflection. Effectively, our scenario is equivalent to a P3P in which an "imaginary" camera $\{\mathbf{C}^*\}$ with a left-handed reference frame lies behind the mirror and observes the true points (not the reflections). To bring (5) into a form similar to (8), we convert the imaginary camera to a right-handed system by pre-multiplying with a reflection across the y-axis (although any axis can be chosen):

$$\begin{aligned} \begin{bmatrix} {}^{\check{C}}\mathbf{p}' \\ 1 \end{bmatrix} &= \begin{bmatrix} \left(\mathbf{I}_3 - 2\mathbf{e}_2\mathbf{e}_2^{\mathrm{T}}\right) & \mathbf{0}_{3\times1} \\ \mathbf{0}_{1\times3} & 1 \end{bmatrix} \begin{bmatrix} \left(\mathbf{I}_3 - 2\,{}^C\mathbf{n}{}^C\mathbf{n}^{\mathrm{T}}\right) & 2d\,{}^C\mathbf{n} \\ \mathbf{0}_{1\times3} & 1 \end{bmatrix} \begin{bmatrix} {}^C_B\mathbf{R} & {}^C\mathbf{p}_B \\ \mathbf{0}_{1\times3} & 1 \end{bmatrix} \begin{bmatrix} {}^B\mathbf{p} \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} {}^{\check{C}}_B\mathbf{R} & {}^{\check{C}}\mathbf{p}_B \\ \mathbf{0}_{1\times3} & 1 \end{bmatrix} \begin{bmatrix} {}^B\mathbf{p} \\ 1 \end{bmatrix} \end{aligned} \tag{9}$$

where $\mathbf{e}_2 = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^{\mathrm{T}}$, and $\{{}^{\check{C}}_B\mathbf{R}, {}^{\check{C}}\mathbf{p}_B\}$ is the transformation between $\{B\}$ and the right-handed frame $\{\check{C}\}$ of the "imaginary" camera behind the mirror. The origin of $\{\check{C}\}$ coincides with that of $\{C^*\}$, their x- and z-axes are common, and their y-axes lie in opposite directions. Note that this additional reflection can be implemented easily, by simply negating the sign of the y-coordinates of the image measurements.

Applying any P3P solution method to the modified problem in (9), we obtain up to 4 solutions, in general, for the unknown transformation $\{{}^{\check{C}}_B\mathbf{R}, {}^{\check{C}}\mathbf{p}_B\}$. We then reflect each of the solutions back, to obtain:

$$\begin{bmatrix} {}^C\mathbf{p}' \\ 1 \end{bmatrix} = \begin{bmatrix} \left(\mathbf{I}_3 - 2\mathbf{e}_2\mathbf{e}_2^{\mathrm{T}}\right) & \mathbf{0}_{3\times1} \\ \mathbf{0}_{1\times3} & 1 \end{bmatrix} \begin{bmatrix} {}^{\check{C}}_B\mathbf{R} & {}^{\check{C}}\mathbf{p}_B \\ \mathbf{0}_{1\times3} & 1 \end{bmatrix} \begin{bmatrix} {}^B\mathbf{p} \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{b} \\ \mathbf{0}_{1\times3} & 1 \end{bmatrix} \begin{bmatrix} {}^B\mathbf{p} \\ 1 \end{bmatrix} \tag{10}$$

where the pair $\{\mathbf{A}, \mathbf{b}\}$, describes a reflection and a homogeneous transformation. Equating (5) and (10), we observe that:

[2] The indices in this paper are: i for points, j for images, and k for solutions.

$$\mathbf{A} = \left(\mathbf{I}_3 - 2\,{}^{C}\mathbf{n}{}^{C}\mathbf{n}^{\mathrm{T}}\right){}^{C}_{B}\mathbf{R} \tag{11}$$

$$\mathbf{b} = \left(\mathbf{I}_3 - 2\,{}^{C}\mathbf{n}{}^{C}\mathbf{n}^{\mathrm{T}}\right){}^{C}\mathbf{p}_B + 2d\,{}^{C}\mathbf{n}. \tag{12}$$

To summarize, in order to exploit the similarity of our problem to the P3P, we execute the following steps: First, the y-coordinates of the image measurements are negated. Then, the measurements are processed by a P3P algorithm to obtain up to 4 solutions $\{{}^{\check{C}}_{B}\mathbf{R}, {}^{\check{C}}\mathbf{p}_B\}$. Subsequently we employ (10), to obtain up to 4 solutions for $\mathbf{A}$ and $\mathbf{b}$. In the next section, we describe our approach for recovering the unknowns, $\{{}^{C}_{B}\mathbf{R}, {}^{C}\mathbf{p}_B, {}^{C}\mathbf{n}, d\}$, from $\mathbf{A}$ and $\mathbf{b}$ using (11) and (12).

3.3 Solution from 3 points in 3 images

We first examine the number of measurements required for a unique solution. When less than 3 points are observed, regardless of the number of images, there is not enough information to determine the transformation, since 3 non-collinear points are required to define the base frame of reference[3]. From 1 image with 3 points, there are not enough constraints to determine the unknowns [cf. (5) and (6)]. From 2 images with 3 points observed in each, the number of constraints equals the number of unknowns; however, in this case rotations of the 2 mirror planes about the axis of their intersection are unobservable, and thus 2 images are not sufficient [6].

From 3 images with 3 points in each, there are 18 scalar measurements [cf. (6)] and 15 unknowns; 6 from $\{{}^{C}_{B}\mathbf{R}, {}^{C}\mathbf{p}_B\}$, and 3 for each mirror configuration $\{\mathbf{n}_j, d_j\}$, $j = 1\ldots3$ [cf. (5)]. This is an overdetermined system, which is nonlinear in the unknown variables. In what follows, we show how to obtain a solution for this system.

Using P3P as an intermediate step, and momentarily ignoring multiple solutions, we obtain constraints of the form (11), (12) for each of the 3 images: $\{\mathbf{A}_j, \mathbf{b}_j\}$, $j = 1\ldots3$. For each pair of images, $j, j' \in \{1\ldots3\}$, we define the unit vector $\mathbf{m}_{jj'}$, as the perpendicular direction to $\mathbf{n}_j$ and $\mathbf{n}_{j'}$ (i.e., $\mathbf{n}_j^{\mathrm{T}}\mathbf{m}_{jj'} = \mathbf{n}_{j'}^{\mathrm{T}}\mathbf{m}_{jj'} = 0$). Alternatively stated, $\mathbf{m}_{jj'} = \alpha\mathbf{n}_j \times \mathbf{n}_{j'}$, where α is a scaling constant to ensure unit length. Using (11), we obtain:

$$\mathbf{A}_j\mathbf{A}_{j'}^{\mathrm{T}}\mathbf{m}_{jj'} = \left(\mathbf{I}_3 - 2\,{}^{C}\mathbf{n}_j{}^{C}\mathbf{n}_j^{\mathrm{T}}\right)\left(\mathbf{I}_3 - 2\,{}^{C}\mathbf{n}_{j'}{}^{C}\mathbf{n}_{j'}^{\mathrm{T}}\right)\mathbf{m}_{jj'} = \mathbf{m}_{jj'}. \tag{13}$$

Thus, by computing the eigenvector corresponding to the unit eigenvalue of $\mathbf{A}_j\mathbf{A}_{j'}^{\mathrm{T}}$, we determine $\mathbf{m}_{jj'}$ up to sign (it can be shown that $\mathbf{A}_j\mathbf{A}_{j'}^{\mathrm{T}}$ is a special orthogonal matrix with 2 complex conjugate eigenvalues, and 1 eigenvalue equal to 1). Employing the properties of the cross product, we obtain:[4]

[3] In the case of 2 points, or 3 or more collinear points, rotations about the line that the points lie on are not observable.

[4] For the remainder of the paper, we drop the superscript 'C' from $\mathbf{n}_j,\ j = 1\ldots3$.

$$\mathbf{n}_1 = \frac{\mathbf{m}_{13}\times\mathbf{m}_{12}}{||\mathbf{m}_{13}\times\mathbf{m}_{12}||}, \quad \mathbf{n}_2 = \frac{\mathbf{m}_{21}\times\mathbf{m}_{23}}{||\mathbf{m}_{21}\times\mathbf{m}_{23}||}, \quad \mathbf{n}_3 = \frac{\mathbf{m}_{13}\times\mathbf{m}_{23}}{||\mathbf{m}_{13}\times\mathbf{m}_{23}||}. \tag{14}$$

Once we have determined the unit vectors corresponding to the 3 mirror planes, the rotation matrix, ${}^C_B\mathbf{R}$, can be computed independently from 3 sets of equations:

$${}^C_B\mathbf{R}_j = \left(\mathbf{I} - 2\mathbf{n}_j\mathbf{n}_j^{\mathrm{T}}\right)\mathbf{A}_j, \quad j = 1\ldots3. \tag{15}$$

In order to utilize all the available information, and to reduce numerical errors, we seek to compute an “average” ${}^C_B\mathbf{R}$ from these 3 sets of equations. However, employing the arithmetic mean is inappropriate since the property of orthonormality is not maintained. We address this issue with the procedure described in Appendix 1.

Once the rotation, ${}^C_B\mathbf{R}$, and the mirror normal vectors, $\mathbf{n}_j,\ j = 1\ldots3$, are determined, the remaining unknowns $\{{}^C\mathbf{p}_B,\ d_1,\ d_2,\ d_3\}$ appear linearly in the constraint equations [cf. (12)]

$$\begin{bmatrix} \left(\mathbf{I}-2\mathbf{n}_1\mathbf{n}_1^{\mathrm{T}}\right) & 2\mathbf{n}_1 & \mathbf{0}_{3\times1} & \mathbf{0}_{3\times1} \\ \left(\mathbf{I}-2\mathbf{n}_2\mathbf{n}_2^{\mathrm{T}}\right) & \mathbf{0}_{3\times1} & 2\mathbf{n}_2 & \mathbf{0}_{3\times1} \\ \left(\mathbf{I}-2\mathbf{n}_3\mathbf{n}_3^{\mathrm{T}}\right) & \mathbf{0}_{3\times1} & \mathbf{0}_{3\times1} & 2\mathbf{n}_3 \end{bmatrix} \begin{bmatrix} {}^C\mathbf{p}_B \\ d_1 \\ d_2 \\ d_3 \end{bmatrix} = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \mathbf{b}_3 \end{bmatrix} \Leftrightarrow \mathbf{D}\mathbf{x} = \mathbf{c} \tag{16}$$

where $\mathbf{D}$ is a 9×6 known matrix, $\mathbf{c}$ is a 9×1 known vector, and $\mathbf{x}$ is the 6×1 vector of unknowns. The least-squares solution for $\mathbf{x}$ in this linear system is $\mathbf{x} = \mathbf{D}^{\dagger}\mathbf{c}$, where $\mathbf{D}^{\dagger}$ denotes the Moore-Penrose generalized inverse of $\mathbf{D}$. From (14), (15), (23), and (16) the mirror configurations, as well as the camera-to-base transformation are computed.

Up to this point, we assumed that the P3P solution was unique, however, there may be up to 4 solutions per image. Recall that 3 images are required to compute the camera-to-base transformation analytically, hence, there are up to 64 solutions for $\{{}^C_B\mathbf{R}, {}^C\mathbf{p}_B, d_1, d_2, d_3, \mathbf{n}_1, \mathbf{n}_2, \mathbf{n}_3\}$, arising from the $4\times4\times4$ possible combinations of P3P solutions. When the measurements are noiseless, we have observed in simulations that only one of these solutions yields a zero-reprojection error (i.e., satisfies all the constraints exactly). This is because the problem at hand is over-constrained (18 constraints for 15 unknowns), and we expect to have a unique solution. In the presence of pixel noise, none of the solutions will satisfy the measurements exactly, thus, we choose the one with the minimum reprojection error.

Moreover, when $N_c > 3$ images are available, there are $N_s = \binom{N_c}{3}$ analytically computed transformations. However, some of these may be inaccurate as a result of degenerate sets of measurements (e.g., when 3 images are all taken from similar viewing angles). In order to identify the correct solution, we employ spectral clustering to determine the largest set of similar solutions [16]. Specifically, we adopt the unit-quaternion representation of rotation [19], ${}^C\bar{q}_B$, which corresponds to ${}^C_B\mathbf{R}$, and denote each solution as $\{{}^C\bar{q}_B^{(k)}, {}^C\mathbf{p}_B^{(k)}\}$ for $k = 1\ldots N_s$. To perform spectral clustering, we define an affinity matrix, $\mathbf{L}$, in which each element is the Mahalanobis distance between a pair of solutions, indexed by k and k':

$$\mathbf{L}_{kk'} = \begin{bmatrix}\delta\boldsymbol{\theta}_{kk'}^{\mathrm{T}} & \delta\mathbf{p}_{kk'}^{\mathrm{T}}\end{bmatrix} \left[\left(\mathbf{H}_k^{\mathrm{T}}\mathbf{Q}^{-1}\mathbf{H}_k\right)^{-1} + \left(\mathbf{H}_{k'}^{\mathrm{T}}\mathbf{Q}^{-1}\mathbf{H}_{k'}\right)^{-1}\right]^{-1} \begin{bmatrix}\delta\boldsymbol{\theta}_{kk'} \\ \delta\mathbf{p}_{kk'}\end{bmatrix} \tag{17}$$

where $\delta\boldsymbol{\theta}_{kk'}$ is the quaternion error-angle vector between ${}^{C}\bar{q}_{B}^{(k)}$ and ${}^{C}\bar{q}_{B}^{(k')}$ [19], and $\delta\mathbf{p}_{kk'} = {}^{C}\mathbf{p}_{B}^{(k)} - {}^{C}\mathbf{p}_{B}^{(k')}$ is the difference between the translation vectors. The matrices $\mathbf{H}_k$ and $\mathbf{H}_{k'}$ are the measurement Jacobians with respect to the transformation [6], and $\mathbf{Q} = \sigma_\eta^2\mathbf{I}_2$ is the covariance of the pixel noise. We compute the transformation, $\{{}^{C}_{B}\mathbf{R}, {}^{C}\mathbf{p}_{B}\}$, from the largest spectral cluster. The rotation, ${}^{C}_{B}\mathbf{R}$, is determined from (23) using all the quaternions in the cluster (cf. Appendix 1), and the translation, ${}^{C}\mathbf{p}_{B}$, is computed as the arithmetic mean of the translations in the cluster.

3.4 Refining the Solution

Due to the presence of pixel noise, and the fact that noise was not accounted for in the analytical solution, the result of the procedure presented in Sections 3.2-3.3 may be coarse (cf. Section 4). Hence, we employ an MLE to refine our analytically computed estimate. We now present an overview of the MLE for determining the unknown transformation between the camera and base frame [7]. Let the vector of all unknown parameters be denoted by $\mathbf{x}$. This vector comprises the unknown transformation, as well as the parameters $\{{}^{C}\mathbf{n}_j, d_j\}$, $j = 1\ldots N_c$, that describe each mirror configuration:

$$\mathbf{x} = \begin{bmatrix}{}^{C}\mathbf{p}_{B}^{\mathrm{T}} & {}^{C}\bar{q}_{B}^{\mathrm{T}} & {}^{C}\mathbf{n}_1^{\mathrm{T}} & d_1 & \ldots & {}^{C}\mathbf{n}_{N_c}^{\mathrm{T}} & d_{N_c}\end{bmatrix}^{\mathrm{T}}. \tag{18}$$

Assuming Gaussian pixel noise, the likelihood of the measurements is given by:

$$\begin{aligned} L(\mathscr{Z};\mathbf{x}) &= \prod_{i=1}^{N_p}\prod_{j=1}^{N_c} p(\mathbf{z}_{ij};\mathbf{x}) = \prod_{i=1}^{N_p}\prod_{j=1}^{N_c} \frac{1}{2\pi\sigma_\eta^2}\exp\left[-\frac{\left(\mathbf{z}_{ij}-\mathbf{h}\left({}^{C_j}\mathbf{p}_i'\right)\right)^{\mathrm{T}}\left(\mathbf{z}_{ij}-\mathbf{h}\left({}^{C_j}\mathbf{p}_i'\right)\right)}{2\sigma_\eta^2}\right] \\ &= \prod_{i=1}^{N_p}\prod_{j=1}^{N_c} \frac{1}{2\pi\sigma_\eta^2}\exp\left[-\frac{\left(\mathbf{z}_{ij}-\mathbf{h}_{ij}(\mathbf{x})\right)^{\mathrm{T}}\left(\mathbf{z}_{ij}-\mathbf{h}_{ij}(\mathbf{x})\right)}{2\sigma_\eta^2}\right] \end{aligned}$$

where the dependence on $\mathbf{x}$ is explicitly shown [cf. (5), (6)], and N_p is the total number of points observed in each of the N_c images. Maximizing the likelihood is equivalent to minimizing its negative logarithm, or minimizing the cost function:

$$c(\mathbf{x}) = \sum_{i=1}^{N_p}\sum_{j=1}^{N_c}(\mathbf{z}_{ij} - \mathbf{h}_{ij}(\mathbf{x}))^{\mathrm{T}}(\mathbf{z}_{ij} - \mathbf{h}_{ij}(\mathbf{x})). \tag{19}$$

We solve this nonlinear least-squares problem with Gauss-Newton iterative minimization to estimate $\mathbf{x}$. At each iteration, indexed by ℓ, the estimate is changed by

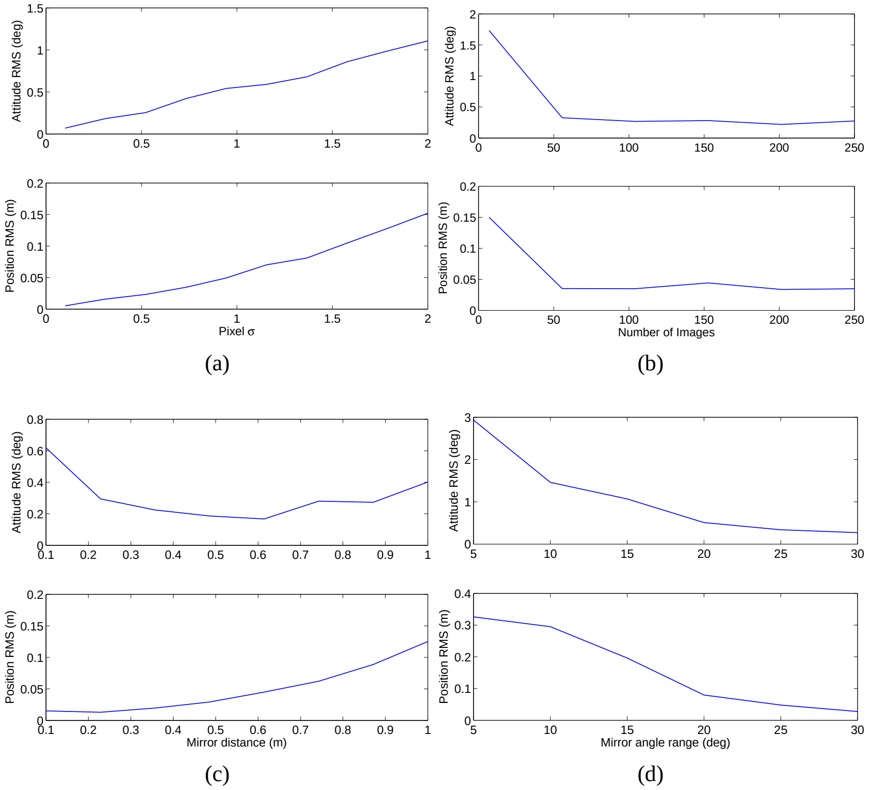


Fig. 2: Average RMS error over 10 trials for attitude and position plotted versus: (a) pixel noise, (b) number of images, (c) mirror distance, and (d) range of mirror rotation.

$$\delta \mathbf{x}^{(\ell)} = \left(\sum_{i,j} \mathbf{J}_{ij}^{(\ell)\mathrm{T}} \mathbf{J}_{ij}^{(\ell)} \right)^{-1} \left(\sum_{i,j} \mathbf{J}_{ij}^{(\ell)\mathrm{T}} \left(\mathbf{z}_{ij} - \mathbf{h}_{ij}(\mathbf{x}^{(\ell)}) \right) \right)$$

where $\mathbf{J}_{ij}^{(\ell)}$ is the Jacobian of $\mathbf{h}_{ij}$ with respect to $\mathbf{x}$, evaluated at the current iterate, $\mathbf{x}^{(\ell)}$. The analytically computed solution from Sections 3.2-3.3 is utilized as the initial iterate, $\mathbf{x}^{(0)}$. Since the MLE is not the main contribution of this work, we limit our discussion here, but refer the reader to [7] for more details.

4 Simulations

In this section, we study the accuracy of the analytically computed camera-to-base transformation (cf. Sections 3.2-3.3). In particular, we investigate how the accuracy is affected by the following parameters: (i) pixel noise, (ii) number of images,

(iii) distance from camera to mirror, and (iv) range of the mirror's angular motion. We consider a "standard" case, in which 3 points placed at the corners of a right triangle with sides measuring $20 \times 20 \times 20\sqrt{2}$ cm are observed in 200 images, while a mirror placed at a distance of 0.5 m is rotated by 30^{o} in two directions. We vary each of the aforementioned parameters individually to examine its effect on the solution accuracy. In Fig. 2, we plot the average RMS error for the position and attitude, computed over 10 trials. Some key observations are:

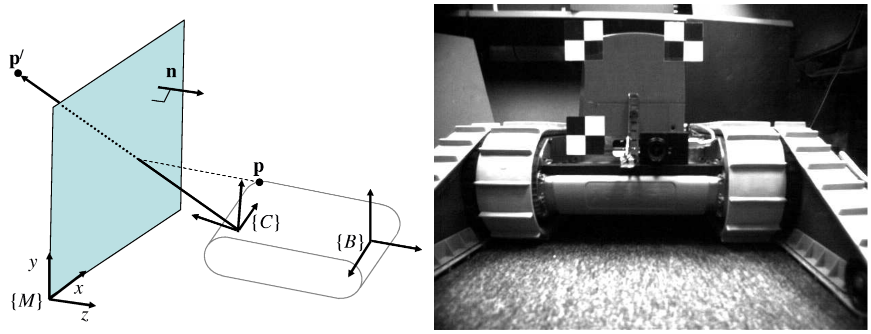
- Increasing the camera's pixel noise decreases the accuracy of the computed solution. When the camera measurements become substantially noisy, e.g., $\sigma = 2$ pixels, the average RMS error is 1^{o} in attitude and 15 cm in position.
- Increasing the number of images results in higher accuracy. However, the improvement follows the "law of diminishing returns," i.e., when a large number of images is already available, the impact of recording more observations is smaller.
- Changing the distance from the mirror to the camera has a significant effect on the position accuracy. When the mirror is at a distance of 1 m, the average RMS error for position is approximately 13 cm. The magnitude of this error suggests that the mirror distance should be kept small. Additionally, it highlights the need to refine our analytically computed transformation with an MLE. As we show in [7], the accuracy of the MLE is approximately 5 times better in attitude, and 10 times better in position compared to the analytical solution.
- Increasing the range of the mirror's angular motion results in improved accuracy. The effect is significant and every effort should be made to move the mirror in the widest range of motion allowed by the camera's field of view.

As a final remark, we note that using the analytical solution as an initial guess for the MLE enables the latter to converge to the correct minimum 100% of the time. On average, fewer iterations were required (approx. 7) when compared to using a naïve initial guess (approx. 18). This shows that a precise analytical solution improves the speed and robustness of the overall estimation process.

5 Experiments

The method described in the preceding sections was employed for computing the transformation between a camera and a base frame attached on the robot-body. For this purpose, 3 fiducial points were placed in known positions on the robot as shown in Fig. 3b. The origin of $\{B\}$ coincides with the top-left fiducial point; both $\{B\}$ and $\{C\}$ are right-handed systems with the axes of $\{B\}$ approximately aligned with those of $\{C\}$. These points were tracked using the KLT algorithm [18] in 1000 images, recorded by a Firewire camera with resolution of 1024×768 pixels.

A planar mirror was maneuvered in different spatial configurations (rotating about two axes), and in distances varying between 30 and 50 cm from the camera, in order to generate a wide range of views. All the measurements were processed to compute the transformation analytically: ${}^{C}\mathbf{p}_{B} = \begin{bmatrix}-14.13 & -10.25 & -13.89\end{bmatrix}^{\mathrm{T}}$ cm,



(a) (b)

Fig. 3: (a) Observation of a point on the robot reflected in the mirror, and (b) an image with 3 fiducial points, captured during experimentation.

and ${}^{C}\bar{q}_B = \begin{bmatrix}-0.0401 & -0.0017 & -0.0145 & 0.9991\end{bmatrix}^{\mathrm{T}}$. This initial solution was refined using the MLE described in Section 3.4, to obtain a better estimate for the transformation between the two frames of interest. The Gauss-Newton iterative minimization converged after 8 iterations, to the following solution for the transformation: ${}^{C}\mathbf{p}_B = \begin{bmatrix}-14.80 & -15.96 & -14.95\end{bmatrix}^{\mathrm{T}}$ cm, and ${}^{C}\bar{q}_B = \begin{bmatrix}0.0045 & 0.0774 & 0.0389 & 0.9962\end{bmatrix}^{\mathrm{T}}$. The corresponding 3σ uncertainty bounds are $\begin{bmatrix}1.1 & 1.6 & 5.0\end{bmatrix}$ mm for the position, and $\begin{bmatrix}0.2419 & 0.2313 & 0.0665\end{bmatrix}$ degrees for the orientation estimates. We point out that the estimates agree with our best guess from manual measurement. We believe that the attained accuracy (given by the 3σ bounds from the MLE) is sufficiently high for most practical applications.

6 Conclusions and Future Work

In this paper, we propose a method for computing the 6 d.o.f. transformation between a camera and a base frame of reference. A mirror is maneuvered in front of the camera, to provide observations of known points from different viewing angles and distances. These measurements are utilized to analytically compute the camera-to-base transformation, and the solution is refined using a maximum-likelihood estimator, which produces estimates for the camera-to-base transformation, as well as for the mirror configuration in each image. The approach was validated both in simulation and experimentally. One of the key advantages of the proposed method is its ease of use; it only requires a mirror, and it provides a solution with as little as 3 points viewed in 3 images. When more information is available, it can be incorporated to produce a more accurate estimate of the transformation.

In our future work, we will investigate the feasibility of mirror-based robot-body 3D reconstruction which we briefly discuss in Appendix 2. Furthermore, we plan to extend this method to the case where the coordinates of the points in the base frame are not known *a priori*, but are estimated along with the camera-to-base transformation and the mirror configurations.

Acknowledgements This work was supported by the University of Minnesota (DTC), and the National Science Foundation (EIA-0324864, IIS-0643680, IIS-0811946). Anastasios Mourikis was partially supported by the UMN Doctoral Dissertation Fellowship, and by the University of California, Riverside. Joel Hesch was supported by the NIH Neuro-physical-computational Sciences Fellowship.

Appendix 1

In this section, we describe the procedure employed for computing an "average rotation," given N_q rotation estimates $\bar{q}_j,\ j = 1 \ldots N_q$. We adopt the quaternion notation from [19] and denote the quaternion of rotation arising from the jth set of equations as $\bar{q}_j$, which corresponds to ${}^C_B\mathbf{R}_j$ [cf. (15)]. Assuming that $\bar{q}$ is the optimal estimate, and employing the small error-angle approximation, we write the following expression for the error in each $\bar{q}_j$:

$$\bar{q}_j \otimes \bar{q}^{-1} \simeq \begin{bmatrix} \hat{\mathbf{k}}_j \frac{\delta\theta_j}{2} \\ 1 \end{bmatrix}, \qquad j = 1 \ldots N_q \tag{20}$$

where $\otimes$ denotes quaternion multiplication, $\hat{\mathbf{k}}_j$ is the unit-vector axis of rotation, and $\delta\theta_j$ is the error angle between the two quaternions. Rewriting this last expression as a matrix-vector multiplication [19], yields

$$\mathscr{L}(\bar{q}_j)\,\bar{q}^{-1} = \begin{bmatrix} \hat{\mathbf{k}}_j \frac{\delta\theta_j}{2} \\ 1 \end{bmatrix}, \qquad j = 1 \ldots N_q \tag{21}$$

where $\mathscr{L}(\bar{q}_j)$, is the left-side quaternion multiplication matrix parameterized by $\bar{q}_j$. Projecting this relation, to keep only the error components, we obtain:

$$\mathbf{P}\mathscr{L}(\bar{q}_j)\,\bar{q}^{-1} = \hat{\mathbf{k}}_j \frac{\delta\theta_j}{2}, \qquad j = 1 \ldots N_q \tag{22}$$

where $\mathbf{P} = \begin{bmatrix} \mathbf{I}_3 & \mathbf{0}_{3\times 1} \end{bmatrix}$. Stacking these relations, we have

$$\begin{bmatrix} \mathbf{P}\mathscr{L}(\bar{q}_1) \\ \vdots \\ \mathbf{P}\mathscr{L}(\bar{q}_{N_q}) \end{bmatrix} \bar{q}^{-1} = \frac{1}{2} \begin{bmatrix} \hat{\mathbf{k}}_1 \frac{\delta\theta_1}{2} \\ \vdots \\ \hat{\mathbf{k}}_{N_q} \frac{\delta\theta_{N_q}}{2} \end{bmatrix}. \tag{23}$$

Our goal is to find the $\bar{q}^{-1}$ that minimizes the norm of the right-hand side. This occurs when $\bar{q}^{-1} = \mathbf{v}(\sigma_{min})$, i.e., we select $\bar{q}^{-1}$ to be the right singular vector corresponding to the minimum singular value of the $3N_q \times 4$ matrix multiplying $\bar{q}^{-1}$ in (23). After finding $\bar{q}^{-1}$ by SVD, we compute the optimal estimate for the rotational matrix ${}^C_B\mathbf{R} = \mathbf{R}(\bar{q})$, which is the rotational matrix parameterized by $\bar{q}$.

Appendix 2

We turn our attention to mirror-based robot-body 3D reconstruction using mirror reflections. We assume that in addition to the 3 points which are known in the robot-body frame, $\{B\}$, we observe another point, $\mathbf{p}_u$, which is *unknown* in $\{B\}$. From one image, we have [cf. (4)]:

$$\beta\, {}^C\mathbf{p}'_{u_0} = \left(\mathbf{I}_3 - 2\,{}^C\mathbf{n}\,{}^C\mathbf{n}^{\mathrm{T}}\right) {}^C_B\mathbf{R}\,{}^B\mathbf{p}_u + \left(\mathbf{I}_3 - 2\,{}^C\mathbf{n}\,{}^C\mathbf{n}^{\mathrm{T}}\right) {}^C\mathbf{p}_B + 2d\,{}^C\mathbf{n} \tag{24}$$

where β is an unknown scale factor and ${}^C\mathbf{p}'_{u_0}$ is the unit vector along the direction of ${}^C\mathbf{p}'_u$. Pre-multiplying both sides by the reflection matrix yields

$$\beta\, \left(\mathbf{I}_3 - 2\,{}^C\mathbf{n}\,{}^C\mathbf{n}^{\mathrm{T}}\right) {}^C\mathbf{p}'_{u_0} = {}^C_B\mathbf{R}\,{}^B\mathbf{p}_u + {}^C\mathbf{p}_B - 2d\,{}^C\mathbf{n}. \tag{25}$$

We assume that the transformation from $\{B\}$ to $\{C\}$, as well as the mirror configuration have been determined using the method outlined in this paper. Hence, the quantities $\{{}^C_B\mathbf{R}, {}^C\mathbf{p}_B, {}^C\mathbf{n},\ d\}$ are known and ${}^C\mathbf{p}'_{u_0}$ is measured, while the quantities $\{\beta, {}^B\mathbf{p}_u\}$ are unknown. From a single image, there are 3 constraints [cf. (25)] and 4 unknowns; hence, we can constrain ${}^B\mathbf{p}_u$ to lie on a line parameterized by β. If the point is observed in 2 consecutive images, then we will have 6 constraints and 5 unknowns, 3 corresponding to the unknown point's coordinates and 2 to the unknown scale factors. In this case, we expect that ${}^B\mathbf{p}_u$ can be determined uniquely.

This problem is analogous to triangulation of a point from two image views ([5], ch. 12). It is solvable when the origin of the camera frame is different for the two views. This corresponds to the quantity $d\,{}^C\mathbf{n}$ changing. Thus, it suffices to either change the distance to the mirror, or the mirror's orientation with respect to the camera. We expect that the location of every unknown point on the robot-body, which is visible in the mirror reflections, can be determined in the body frame of reference, given that it can be reliably tracked in at least 2 images taken from different views.

References

1. X. Brun and F. Goulette. Modeling and calibration of coupled fish-eye CCD camera and laser range scanner for outdoor environment reconstruction. In *Proc. of the Int. Conf. on 3D Digital Imaging and Modeling*, pages 320–327, Montréal, Canada, Aug. 21–23, 2007.
2. K. Daniilidis. Hand-eye calibration using dual quaternions. *Int. Journal of Robotics Research*, 18(3):286–298, 1999.

3. J. Gluckman and S. K. Nayar. Planar catadioptric stereo: Geometry and calibration. In *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pages 22–28, Ft. Collins, CO, June 23–25, 1999.
4. R. M. Haralick, C.-N. Lee, K. Ottenberg, and M. Nölle. Review and analysis of solutions of the three point perspective pose estimation problem. *Int. Journal of Computer Vision*, 13(3):331–356, Dec. 1994.
5. R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521623049, 2000.
6. J. A. Hesch, A. I. Mourikis, and S. I. Roumeliotis. Camera to robot-body calibration using planar mirror reflections. Technical Report 2008-001, University of Minnesota, Dept. of Comp. Sci. & Eng., MARS Lab, July 2008.
7. J. A. Hesch, A. I. Mourikis, and S. I. Roumeliotis. Determining the camera to robot-body transformation from planar mirror reflections. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pages 3865–3871, Nice, France, Sept. 22–26, 2008.
8. M. Inaba, T. Hara, and H. Inoue. A stereo viewer based on a single camera with view-control mechanisms. In *Proc. of the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pages 1857–1865, Yokohama, Japan, July 26–30, 1993.
9. G. Jang, S. Kim, and I. Kweon. Single camera catadioptic stereo system. In *Proc. of the Workshop on Omnidirectional Vision, Camera Networks and Non-classical Cameras*, Beijing, China, Oct. 21, 2005.
10. K. H. Jang, D. H. Lee, and S. K. Jung. A moving planar mirror based approach for cultural reconstruction. *Computer Animation and Virtual Worlds*, 15(3–4):415–423, July 2004.
11. M. Kanbara, N. Ukita, M. Kidode, and N. Yokoya. 3D scene reconstruction from reflection images in a spherical mirror. In *Proc. of the Int. Conf. on Pattern Recognition*, pages 874–897, Hong Kong, China, Aug. 20–24, 2006.
12. R. K. Kumar, A. Ilie, J.-M. Frahm, and M. Pollefeys. Simple calibration of non-overlapping cameras with a mirror. In *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, Anchorage, AK, June 24–26, 2008.
13. A. Martinelli and R. Siegwart. Observability properties and optimal trajectories for on-line odometry self-calibration. In *Proc. of the IEEE Conf. on Decision and Control*, pages 3065–3070, San Diego, CA, Dec. 13–15, 2006.
14. F. M. Mirzaei and S. I. Roumeliotis. A Kalman filter-based algorithm for IMU-camera calibration: Observability analysis and performance evaluation. *IEEE Trans. on Robotics*, 24(5):1143–1156, 2008.
15. S. K. Nayar. Sphereo: Determining depth using two specular spheres and a single camera. In *Proc. of the SPIE Conf. on Optics, Illumination, and Image Sensing for Machine Vision*, volume 1005, pages 245–254, Nov. 1988.
16. A. Y. Ng, M. I. Jordan, and Y. Weiss. On spectral clustering: Analysis and an algorithm. In *Advances in Neural Information Processing Systems*, volume 2, pages 849–856, British Columbia, Canada, Dec. 3–8, 2002.
17. B. K. Ramsgaard, I. Balslev, and J. Arnspang. Mirror-based trinocular systems in robot-vision. In *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pages 499–502, Barcelona, Spain, Sept. 3–7, 2000.
18. J. Shi and C. Tomasi. Good features to track. In *Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition*, pages 593–600, Washington, DC, June 27–July 2, 1994.
19. N. Trawny and S. I. Roumeliotis. Indirect Kalman filter for 3D attitude estimation. Technical Report 2005-002, University of Minnesota, Dept. of Comp. Sci. & Eng., MARS Lab, Mar. 2005.
20. R. Y. Tsai and R. K. Lenz. A new technique for fully autonomous and efficient 3D robotics hand/eye calibration. *IEEE Trans. on Robotics and Automation*, 5(3):345–358, June 1989.
21. S. Wasielewski and O. Strauss. Calibration of a multi-sensor system laser rangefinder/camera. In *Proc. of the Intelligent Vehicles Symposium*, pages 472–477, Detroit, MI, Sept. 25–26, 1995.
22. A. Würz-Wessel and F. K. Stein. Calibration of a free-form surface mirror in a stereo vision system. In *Proc. of the IEEE Intelligent Vehicle Symposium*, volume 2, pages 471–476, Versailles, France, June 17–21, 2002.